JOURNAL OF MULTIVARIATE ANALYSIS 38, 36–50 (1991)

Rates of Convergence in Multivariate Extreme Value Theory

E. OMEY

EHSAL, Stormstraat 2, 1000 Brussels, Belgium

AND

S. T. RACHEV*

University of California, Santa Barbara, California 93106

Communicated by the Editors

We discuss rates of convergence for the distribution of normalized sample extremes to the appropriate limit distribution. We show that the rate of convergence depends on that of the corresponding dependence functions and that of the marginals. The univariate results are well known by now, so we restrict our attention to dependence functions (Sections 2 and 3). In the final section of the paper we obtain a Berry–Esséen type result for multivariate extremes. © 1991 Academic Press, Inc.

1. INTRODUCTION

Let $\mathbf{X}_n = (X_{n,1}, X_{n,2}, ..., X_{n,k})$, $n = 1, 2, ...,$ denote a sequence of independent random vectors (r.v.'s) in $\mathbb{R}^k$ with a common distribution function (d.f.) F. Define the sample maxima as $\mathbf{M}_n = (M_{n,1}, M_{n,2}, ..., M_{n,k})$, where $M_{n,i} = \max\{X_{m,i}, 1 \leqslant m \leqslant n\}$ for each $i = 1, 2, ..., k$. For many d.f.'s there exist normalizing constants $a_{n,i} > 0$ and $b_{n,i} \in \mathbb{R}$ $(i = 1, 2, ..., k)$ such that

$$\frac{\mathbf{M}_n - \mathbf{b}_n}{\mathbf{a}_n} := \left(\frac{M_{n,1} - b_{n,1}}{a_{n,1}}, ..., \frac{M_{n,k} - b_{n,k}}{a_{n,k}}\right) \xrightarrow{d} \mathbf{Y}, \tag{1.1}$$

where $\mathbf{Y} = (Y_1, ..., Y_k)$ is a r.v. with nondegenerate marginals. The d.f. H of

Received April 10, 1990; revised November 2, 1990.

AMS 1980 subject classifications: primary 60F05; secondary 60E15.

Key words and phrases: multivariate extreme values, uniform rates of convergence, regular variation, dependence functions.

* Partially supported by a grant from the National Science Foundation.

0047-259X/91 $3.00

$\mathbf{Y}$ is a so-called max-extreme value d.f. Its marginal H_i must be one of the three extreme-value types $\varphi_\alpha(x) = \exp(-x^{-\alpha})$ $(x > 0,\ \alpha > 0)$, $\Psi_\alpha(x) = \varphi_\alpha(-x^{-1})$ or $\Lambda(x) = \varphi_1(e^{-x})$. The necessary and sufficient conditions on F for the convergence in (1.1) are well known, see, e.g., [9, 12, 13, 14].

The uniform rate of convergence in the univariate case of (1.1) was considered by Fisher and Tippet [7], Anderson [1], Hall and Wellner [10], Smith [21], Davis [6], Cohen [4, 5], Resnick [19], Balkema and deHaan [2]. For other related results we refer to [9, 11, 16, 17]. The common feature of all estimates obtained by the above authors is that they are precise up to an O-term. We shall only quote the following result from the seminal paper of Smith [21]: suppose g is regularly varying with index $\rho > 1$, and

$$\limsup_{x \to \infty} g(x)\, |F(x) - \varphi_1(x)| < \infty,$$

then

$$\Delta_n := \sup_{x \in \mathbf{R}} |P(M_n/n \leqslant x) - \varphi_1(x)| = O(n/g(n)).$$

Zolotarev [22] and Zolotarev and Rachev [23] used the theory of probability metrics to obtain the exact form of the O-term in the Smith theorem. Using the Bergström [3] convolution method adapted for the maxima of i.i.d. random variables Zolotarev and Rachev [23] obtained the estimate

$$\Delta_n \leqslant C \max(\rho_0, \rho_r, \rho_r^{1/(r-1)}) n^{1-r}, \qquad 1 < r \leqslant 2,$$

where C is an absolute constant depending only on r, $\rho_r := \rho_r(F, \varphi_1) := \sup_{x>0} x^r\, |F(x) - \varphi_1(x)|$ is the Kolmogorov weighted semimetric. Omey and Rachev [15] extended the above estimate for $r > 2$ and considered asymptotic expansions for $P(M_n/n \leqslant x)$, see further Lemma 1.4.

The present paper deals with the (uniform) rate of convergence in (1.1) and we extend the univariate results of [15, 23] to the multivariate case. In order to formulate our results, recall the following definition of a dependence function, due to Fréchet [8].

Definition 1.1. Let $F(\mathbf{x})$ be a k-dimensional d.f. with marginals $F_1, F_2, ..., F_k$ and let $D(\mathbf{x})$ be a k-dimensional d.f. on the unit cube $[0, 1]^k$ and with uniformly distributed marginals. The function D (or D_F) is called a dependence function if

$$F(\mathbf{x}) = D(F_1(x_1), ..., F_k(x_k)) \qquad \text{for all} \quad \mathbf{x} \in \mathbb{R}^k.$$

If F has continuous marginals, then F has a unique dependence function. The following classical result motivates the use of dependence functions in the present context.

THEOREM 1.2 [9, Theorem 4.2.3]. *Suppose* $\mathbf{Y}$ *is a nondegenerate r.v. with d.f. H. Then* (1.1) *holds if and only if each marginal in* (1.1) *converges* (*in distribution*) *to a nondegenerate r.v. and for each* $\mathbf{u} \in [0, 1]^k$,

$$\lim_{n \to \infty} D^n(u_1^{1/n}, u_2^{1/n}, ..., u_k^{1/u}) = D_H(u_1, ..., u_k), \tag{1.2}$$

where D denotes a dependence function of F and D_H *the dependence function of H. Moreover,* D_H *satisfies* $D_H^s(u_1^{1/s}, ..., u_k^{1/s}) = D_H(u_1, ..., u_k)$ *for all* $s > 0$.

In order to investigate the rate of convergence in (1.1) wesplit the problem into two independent parts. Note that, first, the rate of convergence (1.1) depends on that of the convergence of all marginals. Second, it depends on the rate of convergence in (1.2). To see that we prove the following estimates in terms of the uniform metric $\rho(\mathbf{U}, \mathbf{V}) := \sup_{\mathbf{x} \in \mathbb{R}^k} |F_{\mathbf{U}}(\mathbf{x}) - F_{\mathbf{V}}(\mathbf{x})|$, where $F_{\mathbf{U}}$ is the d.f. of $\mathbf{U}$.

LEMMA 1.3.

$$\rho\left(\frac{\mathbf{M}_n - \mathbf{b}_n}{\mathbf{a}_n}, \mathbf{Y}\right) \leqslant \sum_{i=1}^{k} \rho\left(\frac{M_{n,i} - b_{n,i}}{a_{n,i}}, Y_i\right) + \sup_{\mathbf{u} \in [0,1]^k} |D^n(u_1^{1/n}, ..., u_k^{1/n}) - D_H(u_1, ..., u_k)|.$$

Proof. Using the definition of D and D_H, we have

$$P\left\{\bigcap_{i=1}^{k} \left\{\frac{M_{n,i} - b_{n,i}}{a_{n,i}} \leqslant x_i\right\}\right\} - H(\mathbf{x})$$
$$= D^n(F_1(a_{n,1}x_1 + b_{n,1}), ..., F_k(a_{n,k}x_k + b_{n,k})) - D_H(H_1(x_1), ..., H_k(x_k)).$$

Hence, by the triangle inequality we have

$$\rho\left(\frac{\mathbf{M}_n - \mathbf{b}_n}{\mathbf{a}_n}, \mathbf{Y}\right) \leqslant \sup_{\mathbf{u} \in [0,1]^k} |D^n(u_1^{1/n}, ..., u_k^{1/n}) - D_H(u_1, ..., u_k)|$$
$$+ \sup_{x \in \mathbb{R}^k} |D_H(F_1^n(a_{n,1}x_1 + b_{n,1}), ..., F_k^n(a_{n,k}x_n + b_{n,k}))$$
$$- D_H(H_1(x_1), ..., H_k(x_k))|.$$

Since D_H has uniform marginals, the second term in the above inequality is not greater than $\sup_{x \in \mathbb{R}^k} \sum_{i=1}^{k} |F_i^n(a_{n,i}x_i + b_{n,i}) - H_i(x_i)|$. This proves the lemma. ∎

The rate of convergence in univariate extreme value theory has been studied extensively. Smith [21] relates uniform rates of convergence to slow variation with remainder. In his paper he considers domain of attraction of $\varphi_\alpha(x)$ and $\psi_\alpha(x)$. Balkema and de Haan [2] obtain uniform rates of convergence to the limit d.f. $\Lambda(x)$. A typical result is the following:

LEMMA 1.4. *Suppose that $X_1, X_2, ..., X_n$ are i.i.d. random variables and Y has d.f. $\varphi_1(x) = \exp(-x^{-1})$ $(x \geqslant 0)$. Let $M_n = \max(X_1, ..., X_n)$.* (i) *For each $r > 1$, $\rho_r(X_1, Y) < \infty$ if and only if* $\lim \text{sum}_{n\to\infty} n^{1-r}\ \rho(n^{-1}M_n, Y) < \infty$, *where $\rho_r(U, V) := \sup_x |x|^r\ |F_U(x) - F_V(x)|$.* (ii) *If $1 < r \leqslant 2$, then there exists a constant $C = C(r) < 682$, such that $\rho(n^{-1}M_n, Y) \leqslant C \max\{\rho(X_1, Y), \rho_r(X_1, Y), [\rho_r(X_1, Y)]^{1/(r-1)}\} n^{1-r}$ for all $n \geqslant 2$.*

Proof. (i) This is [15, Corollary 2.2]. (ii) This is [23, Theorem 2]. ∎

Lemma 1.4 (ii) may be viewed as the analogue of the Berry–Esséen theorem in the extreme-value setting. For corresponding results in the summation-setting we refer to [20] and the references there. The paper is organized as follows. In Section 2 we restrict attention to rates of convergence for dependence functions. In Section 3 some examples are given. In Section 4 we generalize Lemma 1.4 to dimension $k \geqslant 2$ and to all values of $r > 1$.

2. Rates of Convergence for Dependence Function

From Lemma 1.3 we see that it will be useful to estimate the rate of convergence to zero of

$$r_n := \sup_{\mathbf{u} \in [0,1]^k} |D^n(u_1^{1/n}, ..., u_k^{1/n}) - D_H(u_1, ..., u_k)|.$$

In order to formulate our results we require r_n in a more suitable form. Recall that the marginals of D_H and D are uniformly distributed. By using the transformation $x = -1/\log u$ for each of the marginals we can rewrite r_n as

$$r_n = \sup_{\mathbf{x} \in [0,\infty)^k} |K^n(nx_1, ..., nx_k) - G(x_1, ..., x_k)|, \tag{2.1}$$

where each of the marginals of K and G has the d.f. $\varphi_1(x) = e^{-1/x}$ $(x \geqslant 0)$ and G is simple max-stable (cf. [12]). In estimating r_n we shall formulate the results in terms of the metric ρ_ψ which is defined below. Let $\mathbf{V}$ and $\mathbf{W}$ denote two nonnegative r.v. in $\mathbb{R}^k$ and let $\psi: [0, \infty)^k \to [0, \infty)$ denote a continuous function, increasing to infinity in each component. We define the weighted Kolomogorov metric,

$$\rho_\psi(\mathbf{V}, \mathbf{W}) := \sup_{\mathbf{x} \in [0, \infty)^k} \psi(\mathbf{x}) \, |F_\mathbf{V}(\mathbf{x}) - F_\mathbf{W}(\mathbf{x})|.$$

Let $\varphi(x) = \psi(x, x, ..., x)$, $\|\mathbf{x}\| = \min(x_1, ..., x_k)$ and $\rho_\varphi(\mathbf{V}, \mathbf{W}) = \sup_{x \in [0,1]^k} \varphi(\|x\|) \, |F_\mathbf{V}(x) - F_\mathbf{W}(x)| \leqslant \rho_\psi(\mathbf{V}, \mathbf{W})$. We shall often refer to the following inequalities. For $a, b > 0$

$$n\,|a-b| \min(a^{n-1}, b^{n-1}) \leqslant |a^n - b^n| \leqslant n(a-b) \max(a^{n-1}, b^{n-1}). \quad (2.2)$$

If F is a d.f. with marginals $F_i(x) = \varphi_1(x)$, then automatically

$$F(\mathbf{x}) \leqslant \varphi_1(\|\mathbf{x}\|). \quad (2.3)$$

For the function $\varphi(x)$ defined above denote $g(a) := \sup_{x \geqslant 0} (\varphi_1(xa)/\varphi(x))$, $a \geqslant 0$. If $g(a) < \infty$ for all $a > 0$ we define $R(n) := ng(1/(n-1))$ $(n \geqslant 2)$. Finally, let $\mathbf{V}$ be a r.v. with d.f. K and $\mathbf{W}$ a r.v. with d.f. G (cf. (2.1)). In the next theorem we give a sharp estimate for r_n in (2.1).

THEOREM 2.1. (i) *Assume* $\rho_\psi(\mathbf{V}, \mathbf{W}) < \infty$, *then for each* $n \geqslant 2$ *we have* $r_n \leqslant R(n)\, \rho_\psi(\mathbf{V}, \mathbf{W})$. (ii) *If* $\rho_\varphi(\mathbf{V}, \mathbf{W}) < \infty$, *then for each* $n \geqslant 2$, *we have* $r_n \leqslant R(n)\, \rho_\varphi(\mathbf{V}, \mathbf{W})$.

Proof. We only prove (i), (ii) can be proved in a similar way. Using (2.2) and $G(\mathbf{x}) = G^n(n\mathbf{x})$ we have $|K^n(n\mathbf{x}) - G(\mathbf{x})| \leqslant n\,|K(n\mathbf{x}) - G(n\mathbf{x})| \max(K^{n-1}(n\mathbf{x}), G^{n-1}(n\mathbf{x}))$. Using (2.3) for K and G we have

$$|K^n(n\mathbf{x}) - G(\mathbf{x})| \leqslant n\,|K(n\mathbf{x}) - G(n\mathbf{x})| \; \varphi_1^{n-1}(n\,\|\mathbf{x}\|). \quad (2.4)$$

Hence

$$|K^n(n\mathbf{x}) - G(\mathbf{x})| \leqslant \psi(n\mathbf{x})\,|K(n\mathbf{x}) - G(n\mathbf{x})|\, n \, \frac{\varphi_1(n\,\|\mathbf{x}\|)}{(n-1)} \Big/ \psi(n\mathbf{x})$$

$$\leqslant \rho_\psi(\mathbf{V}, \mathbf{W})\, n \, \frac{\varphi_1(n\,\|\mathbf{x}\|)}{n-1} \Big/ \varphi(n\,\|\mathbf{x}\|) \leqslant \rho_\psi(V, W)\, ng\left(\frac{1}{n-1}\right).$$

This proves the theorem. ∎

If, for example, $\varphi(x) = x^r$, $(r \geqslant 1)$, then $g(a) = a^r B(r)$, where $B(r) = (r/e)^r$. If, more generally, $\varphi(x) \in RV_r$, $r \geqslant 1$, we have $g(1/n) \sim B(r)(1/\varphi(n))$ as $n \to \infty$ [15]. Obviously, Theorem 2.1 is only useful if $\lim_{x \to \infty} (\varphi(x)/x) = \infty$.

The next theorem shows that the conditions imposed in Theorem 2.1 are almost necessary.

THEOREM 2.2. *Assume that* $\varphi \in RV_r$, $r \geqslant 1$ *and that* $\lim_{x \to \infty} (\varphi(x)/x) = \infty$. *Then*

(i) $\rho_\varphi(\mathbf{V}, \mathbf{W}) < \infty$ *holds if and only if* $\limsup_{n \to \infty} (\varphi(n)/n) r_n < \infty$, *and*

(ii) $\lim_{\|x\| \to \infty} \varphi(\|\mathbf{x}\|) \mid K(\mathbf{x}) - G(\mathbf{x})| = 0$ *if and only if* $\lim_{n \to \infty} (\varphi(n)/n) r_n = 0$.

Proof. (i) If $\rho_\varphi(\mathbf{V}, \mathbf{W}) < \infty$, the result follows as in Theorem 2.1. To prove the "if" part we use (2.2) to obtain $n|K(\mathbf{x}) - G(\mathbf{x})| \min(K^{n-1}(\mathbf{x}), G^{n-1}(\mathbf{x})) \leqslant r_n$. Now suppose $\|\mathbf{x}\| > N$; choose n such that $n \leqslant \|\mathbf{x}\| \leqslant n+1$. Obviously $K^{n-1}(\mathbf{x}) \geqslant K^{n-1}(\|\mathbf{x}\| \mathbf{1}) \geqslant K^{n-1}(n\mathbf{1})$ and $G^{n-1}(\mathbf{x}) \geqslant G^{n-1}(n\mathbf{1})$, where $\mathbf{1} = (1, ..., 1) \in \mathbb{R}^k$. Since $\lim_{n \to \infty} r_n = 0$ it follows that $\lim_{n \to \infty} K^{n-1}(n\mathbf{1}) = \lim_{n \to \infty} G^{n-1}(n\mathbf{1}) = G(\mathbf{1}) > 0$. Hence

$$\varphi(\|\mathbf{x}\|)\, |K(\mathbf{x}) - G(\mathbf{x})| \leqslant \frac{\varphi(n+1)}{\varphi(n)} \cdot \frac{\varphi(n) r_n}{n} \cdot \frac{1}{\min(K^{n-1}(n\mathbf{1}), G^{n-1}(n\mathbf{1}))}.$$

Since $\varphi \in RV_r$, we have $\varphi(n+1) \sim \varphi(n)$ $(n \to \infty)$. Hence $\sup_{\|x\| > N} \varphi(\|\mathbf{x}\|)$ $|K(\mathbf{x}) - G(\mathbf{x})| < \infty$ and consequently $\rho_\varphi(\mathbf{V}, \mathbf{W}) < \infty$.

(ii) If $\lim_{n \to \infty} (\varphi(n)/n) r_n = 0$ it follows as in the proof of part (i) that $\lim_{\|x\| \to \infty} \varphi(\|\mathbf{x}\|)\, |K(\mathbf{x}) - G(\mathbf{x})| = 0$. To prove the "only if" part, choose N such that $\varphi(\|\mathbf{x}\|)\, |K(\mathbf{x}) - G(\mathbf{x})| < \varepsilon$ for all $\|\mathbf{x}\| > N$. Now proceed as in the proof of Theorem 2.1. If $n\|\mathbf{x}\| > N$, from (2.4) we obtain

$$|K^n(n\mathbf{x}) - G(\mathbf{x})| \leqslant \varepsilon\, ng\left(\frac{1}{n-1}\right). \tag{2.5}$$

On the other hand, if $n\|\mathbf{x}\| \leqslant A$ we use (2.3) to see that $|K^n(n\mathbf{x}) - G^n(n\mathbf{x})| \leqslant 2\varphi_1^{n-1}(\|\mathbf{x}\| n) \leqslant 2\varphi_1(A/(n-1))$. Combining the two estimates we find that

$$\limsup_{n \to \infty} \frac{\varphi(n)}{n} r_n \leqslant \varepsilon\, B(r) + \limsup_{n \to \infty} 2\left(\varphi_1\left(\frac{A}{n-1}\right)\right)(\varphi(n)/n).$$

Since $\varphi(x) \in RV_r$ and since $\varphi_1(1/n)$ tends to zero exponentially fast we obtain $\limsup_{n \to \infty} (\varphi(n)/n) r_n \leqslant \varepsilon\, B(r)$. Now let $\varepsilon \downarrow 0$ to obtain the desired result. ∎

Remark. Theorem 2.1 is also applicable for 0-regularly varying functions φ (i.e., functions for which $\limsup_{t \to \infty} (\varphi(xt)/\varphi(t)) < \infty$ for all $x > 0$). Also exponential types of functions $\varphi(x)$ are allowed. If, for example, $\varphi(x)=x^r \exp(cx^s)$ $(s, c>0)$ then it can be proved that $g(1/n)$ is of the form $n^\alpha \exp(-n^\beta)$ $(0<\beta<1)$, [15, p. 604].

3. Examples

In this section we illustrate the previous results for some well known bivariate d.f.

3.1. *Morgenstern-type distributions.* Here $F(x, y)=F_1(x)\,F_2(y)\,(1+\alpha(1-F_1(x))\,(1-F_2(y)))$, where F_1 and F_2 are (nondegenerate) d.f. in $\mathbb{R}$ and $0<\alpha<1$. Obviously $D_F(u, v)=uv(1+\alpha(1-u)(1-v))$ and $D_H(u, v)=uv$. Moreover, $|D(u, v)-uv|\leqslant\alpha(1-u)(1-v)$. An appeal to Theorem 2.1 shows that $\psi(x, y)=xy$, $\rho_\psi(\mathbf{V}, \mathbf{W})\leqslant\alpha$. Since $\varphi(x)=x^2$ it follows that $r_n\leqslant n^{-1}B(2)\,\rho_\psi(\mathbf{V}, \mathbf{W})$.

3.2. *Marshall-Olkin distribution.* Let $F(x, y)=1-e^{-x}-e^{-y}+e^{-x-y-\lambda\max(x, y)}$ $(\lambda>0)$. It is easily seen that $D(u, v)=u+v-1+(1-u)(1-v)(1-\max(u, v))^\lambda$. Again $D_H(u, v)=uv$ and $\rho_\psi(\mathbf{V}, \mathbf{W})<\infty$, where $\psi(x, y)=xy$. It follows that $r_n\leqslant n^{-1}B(2)\,\rho_\psi(\mathbf{V}, \mathbf{W})$.

3.3. Let (X, Z) denote a random vector with $Z=-X$. Then $F(x, z)=P\{X\leqslant x,\ Z\leqslant z\}=F_X(x)-F_X(-z)$ and $F_1(x)=F_X(x)$, $F_2(z)=1-F_X(-z)$. Hence $F(x, y)=F_1(x)+F_2(y)-1$ so that $D(u, v)=u+v-1$. It easily follows that also here $D_H(u, v)=uv$ and that $r_n\leqslant n^{-1}B(2)\,\rho_\psi(\mathbf{V}, \mathbf{W})$ wih $\psi(x, y)=xy$.

4. Berry–Esséen-Type of Result for the Max-Scheme

As we announced in the Introduction, we prove a general Berry–Esséen-type result for the max-scheme, hereby generalizing Lemma 1.4(ii) to dimension $k\geqslant 2$ and all $r>1$. Let $\mathbf{Y}$ denote a r.v. in $\mathbb{R}^k$ and supose the d.f. H of $\mathbf{Y}$ is simple stable, i.e., each marginal $H_i(x)$ is equal to $\varphi_1(x)=\exp -x^{-1}$ $(x>0)$. In the sequel we frequently use properties of the probability metrics v_r, ρ_r, and L defined below.

Define $v_r(\mathbf{X}', \mathbf{X}'')=\sup_{h>0} h^r\ \rho(\mathbf{X}'\vee h\mathbf{Y},\ \mathbf{X}''\vee h\mathbf{Y})$, where $\mathbf{X}'\vee\mathbf{Y}$ denotes a r.v. with d.f. $F_{\mathbf{X}}(x)\cdot F_{\mathbf{Y}}(x)$. Also define $\rho_r(\mathbf{X}', \mathbf{X}'')=\sup_x \|\mathbf{x}\|^r\,|F_{X'}(\mathbf{x})-F_{X''}(\mathbf{x})|$ where, as before, $\|\mathbf{x}\|=\min(|x_1|, ..., |x_k|)$. Finally, the Lévy metric $L(\mathbf{X}', \mathbf{X}'')$ is defined as usual: $L(\mathbf{X}', \mathbf{X}'')=\inf\{\varepsilon>0\colon F_{\mathbf{X}'}(\mathbf{x}-\varepsilon\mathbf{1})-\varepsilon\leqslant F_{\mathbf{X}'}(\mathbf{x}+\varepsilon\mathbf{1})+\varepsilon\}$, where $\mathbf{1}=(1, 1, ..., 1)\in\mathbb{R}^k$. We summarize some of the properties of these metrics. In the lemma below, $\mathbf{Y}$ and $\mathbf{Y}'$ are i.i.d. simple max-stable r.v. independent of the non-negative r.v.'s $\mathbf{X}'$ and $\mathbf{X}''$.

LEMMA 4.1. (i) *For any* $\delta > 0$, $r > 0$ *we have*

$$\rho(\mathbf{X}' \vee \delta\mathbf{Y}, \mathbf{X}'' \vee \delta\mathbf{Y}) \leqslant \delta^{-r} \nu_r(\mathbf{X}', \mathbf{X}''), \tag{4.1}$$

$$\rho(\mathbf{X}' \vee \delta\mathbf{Y}, \mathbf{X}'' \vee \delta\mathbf{Y}) \leqslant B(r)\, \delta^{-r} \rho_r(\mathbf{X}', \mathbf{X}''), \tag{4.2}$$

$$\nu_r(\mathbf{X}', \mathbf{X}'') \leqslant B(r)\, \rho_r(\mathbf{X}', \mathbf{X}''), \tag{4.3}$$

where $B(r) = (r/e)^r$.

(ii) *For any* $\mathbf{X}$ *we have* $\rho(\mathbf{X}, \mathbf{Y}) \leqslant (1 + kB(2))\, L(\mathbf{X}, \mathbf{Y})$.

(iii) *For any* $\delta > 0$ *we have* $L(\mathbf{X}', \mathbf{Y}'') \leqslant \rho(\mathbf{X}' \vee \delta\mathbf{Y}, \mathbf{X}'' \vee \delta\mathbf{Y}) + k\delta$.

(iv) *For any* $\delta > 0$, $\rho(\mathbf{X}', \mathbf{Y}) \leqslant c_1 \rho(\mathbf{X}' \vee \delta\mathbf{Y}, \mathbf{Y}' \vee \delta\mathbf{Y}) + c_2\delta$, *where* $c_1 := 1 + kB(2)$ *and* $c_2 = c_1 k$.

(v) *If* $\mathbf{Z}$ *and* $\mathbf{W}$ *are independent of* $\mathbf{X}'$ *and* $\mathbf{X}''$, *then*

$$\rho(\mathbf{X}' \vee \mathbf{Z}, \mathbf{X}'' \vee \mathbf{Z}) \leqslant \rho(\mathbf{Z}, \mathbf{W})\, \rho(\mathbf{X}', \mathbf{X}'') + \rho(\mathbf{X}' \vee \mathbf{W}, \mathbf{X}'' \vee \mathbf{W}). \tag{4.4}$$

(vi) *For any* $c > 0$, $\nu_r(c\mathbf{X}', c\mathbf{X}'') \leqslant c^r \nu_r(\mathbf{X}', \mathbf{X}'')$, *and* $\rho_r(c\mathbf{X}', c\mathbf{X}'') \leqslant c^r \rho_r(\mathbf{X}', \mathbf{X}'')$.

Remarks. (1) It is not difficult to see that ν_r and ρ_r are max-ideal metric of order r, i.e., $\nu_r(\mathbf{X}' \vee \mathbf{Z}, \mathbf{X}'' \vee \mathbf{Z}) \leqslant \nu_r(\mathbf{X}', \mathbf{X}'')$ and ν_r is homogeneous of order r.

(2) The result of Lemma 4.1(iv) is often called the max-smoothing inequality. The problem of extending it to max-stable sequences is open (cf. [11]).

(3) Expression (v) is a type of "convolution" inequality for the max-scheme. The convolution method goes back to Bergström [3].

(4) Since $\mathbf{Y}$ is concentrated on $\mathbb{R}^k_+$, there is no loss of generality in assuming $\mathbf{X}'$, $\mathbf{X}''$, are nonnegative r.v.'s, see further the statement of Theorem 4.2.

Proof. (i) Inequality (4.1) follows immediately from the definition of ν_r. As to (4.2) note that $|F_{\mathbf{X}' \vee \delta\mathbf{Y}}(\mathbf{x}) - F_{\mathbf{X}'' \vee \delta\mathbf{Y}}(\mathbf{x})| \leqslant |F_{\mathbf{X}'}(\mathbf{x}) - F_{\mathbf{X}''}(\mathbf{x})|\, F_{\mathbf{Y}}(\mathbf{x}/\delta)$. Since $F_{\mathbf{Y}}(\mathbf{x}/\delta) \leqslant F_{Y_1}(\|\mathbf{x}\|/\delta) \leqslant B(r)(\|\mathbf{x}\|/\delta)^r$, we obtain $|F_{\mathbf{X}' \vee \delta\mathbf{Y}}(\mathbf{x}) - F_{\mathbf{X}'' \vee \delta\mathbf{Y}}(\mathbf{x})| \leqslant B(r)\ \delta^{-r} \rho_r(\mathbf{X}', \mathbf{X}'')$ and (4.2) follows. Finally, (4.3) follows from (4.2) and the definition of ν_r.

(ii) Since all marginals of $\mathbf{Y}$ have he same d.f. $\varphi_1(x)$, we have $\rho(\mathbf{X}, \mathbf{Y}) \leqslant (1 + k \sup_x \varphi_1'(x))\ L(\mathbf{X}, \mathbf{Y})$. Since $\varphi_1'(x) \leqslant B(2)$, we obtain the proof of (ii).

(iii) Let $L(\mathbf{X}', \mathbf{X}'') > \gamma$, then there exists $\mathbf{x}_0 \in \mathbb{R}^k_+$ such that $|F_{\mathbf{X}'}(\mathbf{x}) - F_{\mathbf{X}''}(\mathbf{x})| > \gamma$ for $\mathbf{x}_0 \leqslant \mathbf{x} \leqslant \mathbf{x}_0 + \gamma\mathbf{1}$. It follows that $\rho(\mathbf{X}' \vee \delta\mathbf{Y}, \mathbf{X}'' \vee \delta\mathbf{Y}) \geqslant |F_{\mathbf{X}'}(\mathbf{x}_0 + \gamma\mathbf{1}) - F_{\mathbf{X}''}(\mathbf{x}_0 + \gamma\mathbf{1})|\ F_{\delta\mathbf{Y}}(\mathbf{x} + \gamma\mathbf{1}) \geqslant \gamma F_{\delta\mathbf{Y}}(\mathbf{x}_0 + \gamma\mathbf{1}) \geqslant \gamma F_{\mathbf{Y}}(\gamma\mathbf{1}/\delta)$.

Using Fréchet's inequality we obtain $\rho(\mathbf{X}' \vee \delta\mathbf{Y}, \mathbf{X}'' \vee \delta\mathbf{Y}) \geqslant \gamma \max(0, \sum_{i=1}^{k} F_{\mathbf{Y}_i}(\gamma/\delta) - k + 1)$. Since $F_{Y_i}(x) = e^{-1/x} \geqslant 1 - 1/x$, we have $\rho(\mathbf{X}' \vee \delta\mathbf{Y}, \mathbf{X}'' \vee \delta\mathbf{Y}') \geqslant \gamma \max(0, 1 - k\delta/\gamma)$. Hence $\gamma \leqslant \rho(\mathbf{X}' \vee \delta\mathbf{Y}, \mathbf{X}'' \vee \delta\mathbf{Y}) + k\delta$. Now let $\gamma \uparrow L$ to obtain the proof of the desired result.

(iv) Combine parts (ii) and (iii).

(v) For any $\mathbf{x} \in \mathbb{R}^k$ we have $F_{\mathbf{X}' \vee \mathbf{Z}}(\mathbf{x}) - F_{\mathbf{X}'' \vee \mathbf{Z}}(\mathbf{x}) = (F_{\mathbf{X}'}(\mathbf{x}) - F_{\mathbf{X}''}(\mathbf{x}))(F_{\mathbf{Z}}(\mathbf{x}) - F_{\mathbf{W}}(\mathbf{x})) + F_{\mathbf{W}}(\mathbf{x})(F_{\mathbf{X}'}(\mathbf{x}) - F_{\mathbf{X}''}(\mathbf{x}))$ and the result follows.

(vi) Use the definition of ν_r and ρ_r. ∎

We are now ready for the proof of the main theorem of this section.

THEOREM 4.2. *Assume $\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_n$ is a sequence of i.i.d. r.v. in $\mathbb{R}_+^k$ and assume that $\mathbf{Y}$ is simple stable and independent of the $\mathbf{X}_i$. If $\rho_r(\mathbf{X}_1, \mathbf{Y}) < \infty$ for some $r > 1$, then there exists a constant A such that $\rho(n^{-1}\mathbf{M}_n, \mathbf{Y}) \leqslant An^{1-r}$ for all $n \geqslant 1$, where $\mathbf{M}_n = \mathbf{X}_1 \vee \mathbf{X}_2 \vee \cdots \vee \mathbf{X}_n$. Here A depends on r, k, $\nu_r(\mathbf{X}_1, \mathbf{Y})$, $\rho_r(\mathbf{X}_1, \mathbf{Y})$, $\rho(\mathbf{X}_1, \mathbf{Y})$ and will be determined in the course of the proof.*

Proof. Let $\mathbf{Y}, \mathbf{Y}_1, \mathbf{Y}_2, ..., \mathbf{Y}_n$ be i.i.d. simple max-stable random vectors and let $\mathbf{N}_n = \bigvee_{i=1}^{n} \mathbf{Y}_1$. Then $\mathbf{N}_n \overset{d}{=} n\mathbf{Y}$ and using Lemma 4.1(iii) we obtain that

$$\begin{aligned}\rho(n^{-1}\mathbf{M}_n, \mathbf{Y}) &= \rho(n^{-1}\mathbf{M}_n, n^{-1}\mathbf{N}_n)\\ &\leqslant c_1 \rho(n^{-1}\mathbf{M}_n \vee \delta\mathbf{Y}, n^{-1}\mathbf{N}_n \vee \delta\mathbf{Y}) + c_2\delta \qquad (4.5)\end{aligned}$$

for any $\delta > 0$. The next proposition is crucial in the proof of the theorem.

PROPOSITION 1. *For each $\delta > 0$, we have $\rho(n^{-1}\mathbf{M}_n \vee \delta\mathbf{Y}, n^{-1}\mathbf{N}_n \vee \delta\mathbf{Y}) \leqslant \sum_{i=1}^{4} I_i$, where*

$$I_1 := \rho\left(n^{-1} \bigvee_{i=2}^{n} \mathbf{X}_i, n^{-1}\mathbf{Y}_i\right) \rho(n^{-1}\mathbf{X}_1 \vee \delta\mathbf{Y}, n^{-1}\mathbf{Y}_1 \vee \delta\mathbf{Y})$$

$$I_2 := \sum_{j=2}^{m+1} \rho\left(n^{-1} \bigvee_{i=j+1}^{n} \mathbf{X}_i, n^{-1} \bigvee_{i=j+1}^{n} \mathbf{Y}_i\right) \rho\left(n^{-1} \bigvee_{i=1}^{j-1} \mathbf{Y}_i \vee n^{-1}\mathbf{X}_j \vee \delta\mathbf{Y}, n^{-1} \bigvee_{i=1}^{j} \mathbf{Y}_i \vee \delta\mathbf{Y}\right)$$

$$I_3 := \sum_{j=1}^{m+1} \rho\left(n^{-1}\mathbf{X}_j \vee n^{-1} \bigvee_{i=m+2}^{n} \mathbf{Y}_i, n^{-1}\mathbf{Y}_j \vee n^{-1} \bigvee_{i=m+2}^{n} \mathbf{Y}_i\right)$$

$$I_4 := \rho\left(n^{-1} \bigvee_{j=1}^{m+1} \mathbf{Y}_j \vee n^{-1} \bigvee_{j=m+2}^{n} \mathbf{X}_j, n^{-1} \bigvee_{j=1}^{m+1} \mathbf{Y}_j \vee n^{-1} \bigvee_{j=m+2}^{n} \mathbf{Y}_j\right).$$

Here, $1 \leqslant m \leqslant n-1$ and $\bigvee_{j=1}^{0} \mathbf{Z}_j := 0$.

Proof of Proposition 1. The representation

$$F_{\mathbf{X}}^n(n\mathbf{x})-F_{\mathbf{Y}}^n(n\mathbf{x})=\sum_{j=1}^{m+1}(F^{n-j-1}(n\mathbf{x})\,G^{j-1}(n\mathbf{x})-G^j(n\mathbf{x})\,F^{n-j}(x))$$
$$+F^{n-m-1}(n\mathbf{x})\,G^{m+1}(n\mathbf{x})-G^n(n\mathbf{x}),$$

implies that

$$\begin{aligned}
&\rho(n^{-1}\mathbf{M}_n\vee\delta\mathbf{Y},\mathbf{Y}\vee\delta\mathbf{Y}_1)\\
&\quad\leqslant\sum_{j=1}^{m+1}\rho\left(n^{-1}\bigvee_{i=1}^{j-1}\mathbf{Y}_i\vee n^{-1}\sum_{i=j}^{n}\mathbf{X}_i\vee\delta\mathbf{Y},\right.\\
&\qquad\qquad\left.n^{-1}\bigvee_{i=1}^{j}\mathbf{Y}_i\vee n^{-1}\bigvee_{i=j+1}^{n}\mathbf{X}_i\vee\delta\mathbf{Y}\right)\\
&\qquad+\rho\left(n^{-1}\bigvee_{j=1}^{m+1}\mathbf{Y}_j\vee n^{-1}\bigvee_{j=m+2}^{n}\mathbf{X}_j\vee\delta\mathbf{Y},\right.\\
&\qquad\qquad\left.n^{-1}\bigvee_{j=1}^{m+1}\mathbf{Y}_j\vee n^{-1}\bigvee_{j=m+2}^{n}\mathbf{Y}_j\vee\delta\mathbf{Y}\right)=:T_1+T_2.
\end{aligned}$$

Using (4.4) with $\mathbf{Z}:=n^{-1}\bigvee_{i=j+1}^{n}\mathbf{X}_i$ and $\mathbf{W}:=n^{-1}\bigvee_{i=j+1}^{n}\mathbf{Y}_i$ we readily obtain that

$$\begin{aligned}
T_1\leqslant I_1+I_2+\sum_{j=1}^{m+1}\rho&\left(n^{-1}\bigvee_{i=1}^{j-1}\mathbf{Y}_i\vee n^{-1}\mathbf{X}_j\vee n^{-1}\bigvee_{i=j+1}^{n}\mathbf{Y}_i\vee\delta\mathbf{Y},\right.\\
&\left.\quad n^{-1}\bigvee_{j=1}^{n}\vee n^{-1}\bigvee_{i=j+1}^{n}\mathbf{Y}_i\vee\delta\mathbf{Y}\right).
\end{aligned}$$

Since

$$\bigvee_{\substack{i=1\\i\neq j}}^{n}\mathbf{Y}_i=\bigvee_{i=1}^{j-1}\mathbf{Y}_i\vee\bigvee_{\substack{i=j+1\\i\neq j}}^{n}\mathbf{Y}_i=\bigvee_{i=1}^{m+1}\mathbf{Y}_i\vee\bigvee_{i=m+2}^{n}\mathbf{Y}_i,$$

we obtain

$$T_1\leqslant I_1+I_2+\sum_{j=1}^{m+1}\rho\left(n^{-1}\mathbf{X}_j\vee n^{-1}\bigvee_{i=m+2}^{n}\mathbf{Y}_i,n^{-1}\mathbf{Y}_j\vee n^{-1}\bigvee_{i=m+2}^{n}\mathbf{Y}_i\right)$$

or $T_1\leqslant I_1+I_2+I_3$. Finally, it is easy to see that $T_2\leqslant I_4$, which proves the proposition. ∎

In the next proposition we estimate I_3 and I_4. We shall use Proposition 1 with $m=[n/2]$, the integer part of $n/2$.

PROPOSITION 2. *For all* $n \geqslant 3$, $I_3 \leqslant 4^r \frac{2}{3} v_r(\mathbf{X}_1, \mathbf{Y}) n^{1-r}$ *and* $I_4 \leqslant 2^r v_r(\mathbf{X}_1, \mathbf{Y}) n^{1-r}$.

Proof of Proposition 2. (i) Obviously,

$$I_3 = \sum_{j=1}^{m+1} \rho\left(n^{-1}\mathbf{X}_j \vee \frac{n-m-1}{n}\mathbf{Y}, n^{-1}\mathbf{Y}_j \vee \frac{n-m-1}{n}\mathbf{Y}\right).$$

Using (4.1), we obtain

$$I_3 \leqslant \sum_{j=1}^{m+1} \left(\frac{n-m-1}{n}\right)^{-r} v_r(n^{-1}\mathbf{X}_j, n^{-1}\mathbf{Y}_j).$$

Using Lemma 4.1 (vi) and $m = [n/2]$, we arrive at

$$I_3 \leqslant 4^r \tfrac{2}{3} v_r(\mathbf{X}_1, \mathbf{Y}_1) n^{1-r}.$$

(ii) Obviously,

$$I_4 = \rho\left(n^{-1} \bigvee_{j=m+2}^{n} \mathbf{X}_j \vee \frac{m+1}{m} Y, n^{-1} \bigvee_{j=m+2}^{n} \mathbf{Y}_j \vee \frac{m+1}{n}\mathbf{Y}\right).$$

Using (4.1) we obtain

$$I_4 \leqslant \left(\frac{m+1}{n}\right)^{-r} v_r\left(n^{-1} \bigvee_{j=m+2}^{n} \mathbf{X}_j, n^{-1} \bigvee_{j=m+2}^{n} \mathbf{Y}_j\right).$$

Since v_r is max-ideal of order t we obtain $I_4 \leqslant (m+1)^{-r}$ $(n-m-1)$ $v_r(\mathbf{X}_1, \mathbf{Y}_1)$. Hence $I_n \leqslant 2^r n^{1-r} v_r(\mathbf{X}_1, \mathbf{Y}_1)$.

We are now able to prove the theorem in the case where $1 < r \leqslant 2$.

Proof of Theorem 4.2 *in the Case Where* $1 < r \leqslant 2$. We prove the result by induction. Suppose we know that

$$\rho\left(k^{-1} \bigvee_{i=1}^{k} \mathbf{X}_i, k^{-1} \bigvee_{i=1}^{k} \mathbf{Y}_i\right) \leqslant A k^{1-r}, \qquad k = 1, 2, ..., n-1, n \geqslant 3, \quad (4.6)$$

where A will be determined later. We have to prove that (4.6) also holds for $k = n$. In order to do so, we first estimate I_1 and I_2 and we choose $\delta = Bn^{-1}$, where B will be determined later. First consider I_1. Using (4.6) and (4.2) we obtain $I_1 \leqslant A(n-1)^{1-r}$ $\delta^{-r} v_r(n^{-1}\mathbf{X}_1, n^{-1}\mathbf{Y}_1)$. Using Lemma 4.1 (vi) and the choice of δ, we obtain

$$I_1 \leqslant AB^{-r}(\tfrac{3}{2})^r v_r(\mathbf{X}_1, \mathbf{Y}_1) n^{1-r}. \quad (4.7)$$

As to I_2 we have

$$I_2 = \sum_{j=2}^{m+1} \rho\left(n^{-1} \bigvee_{i=j+1}^{n} \mathbf{X}_i, n^{-1} \bigvee_{i=j+1}^{n} \mathbf{Y}_i\right)$$
$$\cdot \rho\left(n^{-1}\mathbf{X}_j \vee \left(\frac{j-1}{n}+\delta\right)\mathbf{Y}, n^{-1}\mathbf{Y}_j \vee \left(\frac{j-1}{n}+\delta\right)\mathbf{Y}\right).$$

Using (4.6) and (4.1) again, we obtain $I_2 \leqslant \sum_{j=2}^{m+1} A(n-j)^{1-r}((j-1)/n+\delta)^{-r}$ $v_r(n^{-1}\mathbf{X}_j, n^{-1}\mathbf{Y}_j)$. Hence, $I_2 \leqslant A4^{r-1}v_r(\mathbf{X}_1, \mathbf{Y}_1)n^{1-r}\sum_{j=1}^{\infty}(j-1+B)^{-r}$ and this implies that

$$I_2 \leqslant \frac{A4^{r-1}v_r(\mathbf{X}_1, \mathbf{Y}_1)}{B^{r-1}(r-1)} n^{1-r}. \tag{4.8}$$

Now combine the estimates (4.7), (4.8) and those of Proposition 2 to obtain $\rho(n^{-1}\mathbf{M}_n \vee \delta\mathbf{Y}, n^{-1}\mathbf{N}_n \vee \delta\mathbf{Y}) \leqslant C_1 n^{1-r}$, where

$$C_1 = \left(\frac{A(3/2)^r}{B^r} + \frac{A4^{r-1}}{B^{r-1}(r-1)} + 4^r\left(\frac{2}{3}\right) + 2^r\right) v_r(\mathbf{X}_1, \mathbf{Y}).$$

Using (4.5) we obtain $\rho(n^{-1}\mathbf{M}_n, \mathbf{Y}) \leqslant c_1 C_1 n^{1-r} + c_2 B n^{-1}$ and hence $(1 < r \leqslant 2)$,

$$\rho(n^{-1}\mathbf{M}_n, \mathbf{Y}) \leqslant (c_1 C_1 + c_2 B 3^{r-2}) n^{1-r}. \tag{4.9}$$

Now we shall choose A and B to ensure (4.6) remains valid for $k=n$. Obviously, $k=1$ and $k=2$ in (4.6) show that $A \geqslant \rho(\mathbf{X}_1, \mathbf{Y}_1)$ and $A \geqslant 2^r\rho(\mathbf{X}_1, \mathbf{Y}_1)$. A close examination of (4.9) shows that $c_1 C + c_2 B 3^{r-2} = \alpha A/B^r + \beta A/B^{r-1} + \gamma + \eta B$, where $\alpha = c_1(\frac{3}{2})^r v_r(\mathbf{X}_1, \mathbf{Y}_1)$, $\beta = c_1(4^{r-1}v_r(\mathbf{X}_1, \mathbf{Y}_1)/(r-1))$, $\gamma = c_1(4^r\frac{2}{3} + 2^r)$ $v_r(\mathbf{X}_1, \mathbf{Y}_1)$, and $\eta = c_2 3^{r-2}$. Now choose B sufficiently large to ensure that $\alpha/B^r - \beta/B^{r-1} \leqslant \frac{1}{2}$. Then choose A large enough so that $A \geqslant \max(2^r\rho(\mathbf{X}_1, \mathbf{Y}_1)\ 2(\gamma + 2B)$. It follows from (4.9) that (4.6) holds for $k=n$.

Remark. Since $1 < r \leqslant 2$ we have $\alpha/B^r + \beta/B^{r-1} \leqslant (\alpha+\beta)/B^{r-1}$. Now choose $B = (2(\alpha+\beta))^{1/(r-1)}$. It is easy to see that $B = \omega(v_r(\mathbf{X}_1, \mathbf{Y}_1))^{1/(r-1)}$, where ω depends only on r and k. As for A we can take

$$A = A^* = \max(2^r\rho(\mathbf{X}_1, \mathbf{Y}_1), \omega_1 v_r(\mathbf{X}_1, \mathbf{Y}_1) + \omega_2 v_r^{1/(r-1)}(\mathbf{X}_1, \mathbf{Y}_1)), \tag{4.10}$$

where ω_1, ω_2 only depend on r and k. It follows that $A \leqslant C \max(\rho(\mathbf{X}_1, \mathbf{Y}_1),$ $v_r(\mathbf{X}_1, \mathbf{Y}_1), v_r^{1/(r-1)}(\mathbf{X}_1, \mathbf{Y}_1))$, where C only depends on r and k.

Proof of Theorem 4.2 *in Case* $r > 2$. If $r > 2$, then $\rho_r(\mathbf{X}_1, \mathbf{Y}) < \infty$ implies that $\rho_2(\mathbf{X}_1, \mathbf{Y}) \leqslant \max(\rho_r(\mathbf{X}_1, \mathbf{Y}), \rho(\mathbf{X}_1, \mathbf{Y})) < \infty$ and (4.6) holds with $r = 2$, all $k \geqslant 1$, with $A = A^*$ (cf. (4.10)):

$$\rho(n^{-1}\mathbf{M}_n, \mathbf{Y}) \leqslant A^* n^{-1}, \qquad n \geqslant 1. \tag{4.11}$$

Choose m so that $1 \leqslant m \leqslant n-1$. Obviously we have

$$\begin{aligned} F^n_{\mathbf{X}_1}(n\mathbf{x}) - F^n_{\mathbf{Y}}(n\mathbf{x}) &= (F^{n-m}_{\mathbf{X}_1}(n\mathbf{x}) - F^{n-m}_{\mathbf{Y}}(n\mathbf{x}))(F^m_{\mathbf{X}_1}(n\mathbf{x}) - F^m_{\mathbf{Y}}(n\mathbf{x})) \\ &\quad + F^{n-m}_{\mathbf{Y}}(n\mathbf{x})(F^m_{\mathbf{X}_1}(n\mathbf{x}) - F^m_{\mathbf{Y}}(n\mathbf{x})) + F^m_{\mathbf{Y}}(n\mathbf{x})(F^{n-m}_{\mathbf{X}_1}(n\mathbf{x}) - F^{n-m}_{\mathbf{Y}}(n\mathbf{x})). \end{aligned}$$

Using $|F^m_{\mathbf{X}_1} - F^m_{\mathbf{Y}}| \leqslant m |F_{\mathbf{X}_1} - F_{\mathbf{Y}}|$ and (cf. (2.3)) $F^k_{\mathbf{Y}}(n\mathbf{x}) = F_{\mathbf{Y}}((n/k)\|\mathbf{x}\|) \leqslant B(r)(n^r/k^r)\|\mathbf{x}\|^r$, we obtain that

$$\begin{aligned} \rho(n^{-1}\mathbf{M}_n, \mathbf{Y}) &\leqslant \rho\left(n^{-1} \bigvee_{i=1}^{n-m} \mathbf{X}_i, n^{-1} \bigvee_{i=1}^{n-m} \mathbf{Y}_i\right) \\ &\quad \cdot \rho\left(n^{-1} \bigvee_{i=1}^{m} \mathbf{X}_i, n^{-1} \bigvee_{i=1}^{m} \mathbf{Y}_1\right) \\ &\quad + B(r)\, \rho_r(\mathbf{X}_1, \mathbf{Y}_1)\left(\frac{m}{(n-m)^r} + \frac{n-m}{m^r}\right). \end{aligned} \tag{4.12}$$

If $m = [n/2]$, $n \geqslant 2$, (4.12) and (4.11) show that

$$\rho(n^{-1}\mathbf{M}_n, \mathbf{Y}) \leqslant \rho(m^{-1}\mathbf{M}_m, \mathbf{Y}) \frac{A^*}{m} + B(r)\, \rho_r((\mathbf{X}_1, \mathbf{Y}_1)\, 3^r n^{1-r}). \tag{4.13}$$

Let $N \geqslant 3$ denote an integer to be determined later. Using (4.11) we have

$$\rho(k^{-1}\mathbf{M}_k, \mathbf{Y}) \leqslant \frac{A^*}{k} \leqslant A^* N^{r-2} k^{1-r}, \qquad k = 1, 2, ..., N. \tag{4.14}$$

Now we prove the result of the theorem by induction. Suppose that

$$\rho(k^{-1}\mathbf{M}_k, \mathbf{Y}) \leqslant K k^{1-r}, \qquad k = 1, 2, ..., n-1, n \geqslant N, \tag{4.15}$$

where K will be determined later. We have to show that (4.15) also holds for $k = n$. Using (4.13) and (4.15),we obtain

$$\begin{aligned} \rho(n^{-1}\mathbf{M}_n, \mathbf{Y}) &\leqslant K m^{1-r} \frac{A^*}{m} + B(r)\, \rho_r(\mathbf{X}_1, \mathbf{Y})\, 3^r n^{1-r} \\ &\leqslant K n^{1-r} \left\{\frac{6^{r-1} A^*}{[N/2]} + \frac{B(r)\, \rho_r(\mathbf{X}_1, \mathbf{Y}) 3^r}{K}\right\} \\ &\leqslant K n^{1-r} \left\{\frac{2 \cdot 6^{r-1} A^*}{N-2} + \frac{B(r)\, \rho_r(\mathbf{X}_1, \mathbf{Y}) 3^r}{K}\right\}. \end{aligned} \tag{4.16}$$

Now choose N sufficiently large to ensure that $N \geqslant 3$ and $4 \cdot 6^{r-1} A^* \leqslant N-2$; i.e., take $N = [4 \cdot 6^{r-1} A^*] + 3$. Next choose K such that (cf. (4.14)) $K \geqslant A^* N^{r-2}$ and (cf. (4.16)) $K \geqslant 2B(r)\ \rho_r(X_1, Y) 3^r$. These choices show that in (4.16) we have $\rho(n^{-1}\mathbf{M}_n, \mathbf{Y}) \leqslant K n^{1-r}$; i.e., (4.15) holds also for $k = n$. ∎

Summarizing our findings for the constant A in Theorem 4.2 we have

COROLLARY 4.3. *In Theorem* 4.2 *we can take*

$$A = \begin{cases} A^* = C \max(\rho(\mathbf{X}_1, \mathbf{Y}), \nu_r(\mathbf{X}_1, \mathbf{Y}), \nu_r^{1/(r-1)}\mathbf{X}_1, \mathbf{Y}) & \text{if } 1 < r \leqslant 2 \\ \quad C \max(A^{*r-1}, \rho(\mathbf{X}_1, \mathbf{Y})) & \text{if } r > 2, \end{cases}$$

where C is a constant which depends only on r and k.

REFERENCES

[1] ANDERSON, C. W. (1971). *Contributions to the Asymptotic Theory of Extreme Values.* Ph.D. thesis, University of London.

[2] BALKEMA, A. A., AND DE HAAN, L. (1990). A convergence rate in extreme value theory. *J. Appl. Probab.*

[3] BERGSTRÖM, H. (1945). On the central limit theorem in $\mathbb{R}^k$. *Skand. Aktuarietidskr.* **28** 106–127.

[4] COHEN, J. P. (1982). The penultimate form of approximation to normal extremes. *Adv. in Appl. Probab.* **14** 324–339.

[5] COHEN, J. P. (1982). Convergence rates for the ultimate and penultimate approximations in extreme value theory. *Adv. in Appl. Probab.* **14** 324–339.

[6] DAVIS, R. (1982). The rate of convergence in distribution of the maxima. *Statist. Neerlandica* **36** 31–35.

[7] FISHER, R. A., AND TIPPET, L. H. C. (1928). Limiting forms of the frequency distributions of largest orsmallest member of a sample. *Proc. Cambridge Philos. Soc.* **24** 180–190.

[8] FRÉCHET, M. (1951/1957/1958). Sur les tableaux de corrélation dont les marges sont donnés. *Ann. Univ. Lyon Sect. A* **14** 53–57; **20** 13–31; **21** 19–32.

[9] GALAMBOS, J. (1987). *The Asymptotic Theory of Extreme Order Statistics,* 2nd ed. Krieger, Malabar, FL.

[10] HALL, W. J., AND WELLNER, J. A. (1979). The rate of convergence in law of the maximum of an exponential sample. *Statist. Neerlandica* **33** 151–154.

[11] DE HAAN, L., AND RACHEV, S. T. (1989). Estimates of the rate of convergence for max-stable processes. *Ann. Probab.* **17** 655–677.

[12] DE HAAN, L., AND RESNICK, S. (1977). Limit theory for multivariate sample extremes. *Z. Wahrsch. Verw. Gebiete* **40** 317–337.

[13] LEADBETTER, M. R., LINDGREN, G., AND ROOTZEN, H. (1980). *Extremes and Related Properties of Random Sequences and Processes.* Springer-Verlag, New York, 1980.

[14] MARSHALL, A. W., AND OLKIN, I. (1983). Domains of attraction of multivariate extreme value distributions. *Ann. Probab.* **11** 168–177.

[15] OMEY, E., AND RACHEV, S. T. (1988). On the rate of convergence in extreme value theory. *Theory Probab. Appl.* **33** 601–607.

[16] RACHEV, S. T. (1991). *Probability Metrics.* Wiley, New York.

[17] REISS, R. D. (1989). *Approximate Distributions of Order Statistics.* Springer-Verlag, New York.

[18] RESNICK, S. I. (1987). *Extreme Values, Regular Variation and Point Processes.* Springer-Verlag, New York.

[19] RESNICK, S. I. (1987). Uniform rates of convergence to extreme-value theory distributions. In *Probability and Statistics: Essay in Honor of Franklin A. Graybill* (J. Srivastava, Ed.), North-Holland, Amsterdam.

[20] SAZONOV, V. V. (1981). *Normal Approximation. Some Recent Advances.* Lecture Notes in Math., Vol. 879. Springer-Verlag, New York, 1981.

[21] SMITH, R. L. (1982) Uniform rates of convergence in extreme value theory. *Adv. in Appl. Probab.* **14** 600–622.

[22] ZOLOTAREV, V. M. (1983). Foreword. In *Stability Problems for Stochastic Models* (V. V. Kalashnikov and V. M. Zolotarev, Eds.). Lecture Notes in Math., Vol. 982. Springer-Verlag, New York.

[23] ZOLOTAREV, V. M., AND RACHEV, S. T. (1985). Rate of convergence in limit theorems for the max-scheme. In *Lecture Notes in Math., Vol. 1155*, pp. 415–442. Springer-Verlag, New York.